## Evaluating the Application of Machine Learning to Control of Advanced Life Support Systems

## Task Leaders

David Kortenkamp, NASA JSC-ER2/Metrica, kortenkamp@jsc.nasa.gov, 281-461-7884, co-PI
Justin Boyan, NASA Ames Research Center, jboyan@mail.arc.nasa.gov, 650-604-3590, co-PI
Terry Tri, NASA JSC-EC, terry.o.tri1@jsc.nasa.gov, 281-483-9234
Jeffrey D. Kowing, NASA JSC-ER2, jeffrey.d.kowing@jsc.nasa.gov
R. Peter Bonasso, Debra Schreckenghost, NASA JSC-ER2/Metrica
Gregory Dorais, NASA ARC, gadorais@ptolemy.arc.nasa.gov
Devika Subramanian, Associate Professor, Rice University, devika@cs.rice.edu

## Product Description

This work will focus on experimentally evaluating different machine learning techniques (reinforcement learning, genetic algorithms explaination-based learning and case-based learning) in the domain of Advanced Life Support Systems (ALSS). The experiments will range from adjusting control parameters to learning new sequences of behaviors. The goal is to determine which machine learning algorithms will best provide the adaptiveness and robustness needed for long-term control of life support systems such as BIO-Plex. A second goal is to determine how machine learning algorithms will interface to autonomous control architectures such as 3T or Remote Agent. Experiments will use actual hardware (provided by JSC-EC) and simulations. The product of this research will be a detailed description, supported by prototypes and experiments, of how machine learning algorithms can integrate with autonomous control architectures to control advanced life support systems.

This is a **new**, **pull** task

## Task Resources and Outline

| | *FY 2000* | *FY 2001* | *FY 2002* |
|---|---|---|---|
| **Budget** | $500K | $500K | $500K |
| **Cost Sharing** | 0 | 0 | 0 |
| **Civil Servants** | 3 | 3 | 3 |
| **Contractors** | 3 | 3 | 3 |
| **Academics** | 1 | 1 | 1 |
| **TRL** | 1 | 2-3 | 3-4 |

**Cost breakout**: $300K JSC, $150K ARC, $50K Rice

## Benefits

The Bioregenerative Planetary Life Support System Test Complex (BIO-Plex) will consist of a complex of 5 chambers at NASA JSC combining biological and physiochemical life support technologies to provide all the air and water, and most of the food for a crew of four on a continuous basis. Initial testing is scheduled for 2001, culminating in a 425-day test starting in 2006. Intelligent, autonomous control has already been identified as a key technology for BIO-Plex, yet there are still substantial research issues that need to be addressed before continuous, autonomous control of BIO-Plex is possible. Among these research issues is machine learning. There will be many unknowns in BIO-Plex and the autonomous control system will have to adapt with a minimum of human intervention. Over the course of a 425-day test many control parameters and sequences will change or be refined and it will be costly to have to use human programmers to keep the autonomous control system up-to-date. Machine learning techniques will allow for lower cost testing (and, eventually, lower cost missions) by reducing both the need for human intervention and the need for reprogramming of the autonomous control system. Machine learning techniques will also result in a more robust control system that can accommodate changes over time. We also expect this research to impact autonomous control of missions to Mars and control of in-situ propellant production systems. As for research benefits, we expect this project to stretch our theoretical understanding of machine learning algorithms, while integration with real-world control systems will enhance our engineering understanding of machine learning algorithms.

## Technical Approach

### Technical abstract

The primary goal of this research is to design an autonomous control system that can keep key parameters of a closed system within pre-determined bounds in a near-optimal way even when there are significant changes in the underlying environment over time or when there are unknown or underspecified aspects to the environment. To use a life support example, this means to keep the levels of gases, water and food within life sustaining ranges even when environmental conditions (e.g., the amount of oxygen produced by plants, the amount of carbon dioxide produced by people, the amount of contamination in the water, the efficiency of biology-based processing machines, etc.) change over time and may not even be known ahead of time. Controlling a system like this will require sophisticated machine learning techniques. In addition, learning needs to occur while the software is controlling the system, so off-line learning techniques that can transfer to on-line control systems will also be a key component of this research.

### Approach

Our approach builds on our experiences in controlling several early advanced life support system tests and the relationships that we have developed with the advanced life support researchers at JSC [Schreckenghost et al 1998a; Schreckenghost et al 1998b]. Our approach also builds on work at Ames [Boyan & Moore, 1998; Moore *et al*, 1998] and work at Rice [Subramanian & Hunter, 1992; Subramanian & Gordon, 1993; Subramanian & Gordon, 1996] in applying machine learning techniques to non-static environments. In bringing together these research efforts we hope to create an integrated solution for adaptively controlling advanced life support systems.

Our technical approach consists of three inter-related tasks:

1. Defining learning tasks in the control of advanced life support systems.
2. Choosing specific learning algorithms applicable to the tasks defined in (1).
3. Designing the interface between machine learning algorithms and autonomous control architectures such as Remote Agent and 3T.

In the first task, we will look at the underlying dynamics of the system we want to control. This will be a broad task that will look at current systems (e.g., the water recovery system (WRS)) that are currently being tested, at near-term systems (e.g., BIO-Plex) that are being built, and at long-term systems (e.g., a Mars base) that are being proposed. The goal of this task is to determine exactly which aspects (or parameters) of these systems are amenable to learning techniques and where machine learning techniques can reduce the need for preprogramming or reprogramming control systems. We expect to look at how machine learning can both adjust low-level parameters (e.g., the setting of a pump) and adjust high-level sequences of actions (e.g., the procedure for restarting a machine). The questions we will need to answer include: what time scales are decisions being made at, what can be learned within a given amount of resources, and how can learning be integrated with the actual controlling of the system. By identifying the learning tasks within ALSS that we want to address and determining which aspects of the system are worth learning we will be ready to start applying learning algorithms to the control system.

In the second task, we will look at a variety of machine learning approaches, including reinforcement learning and genetic algorithms at the real-time control level and explanation-based learning (EBL) and case-based learning (CBL) at the sequencer level. We will evaluate these techniques with respect to the parameters and sequences that need to be learned as defined in Task 1. This investigation will be specific to the ALSS domain. We will be looking at how "rewards" for actions are sampled from the underlying dynamical system and how quickly our learning algorithms need to work compared with the changes to the advanced life support system. This task will give us an understanding of how different learning algorithms will connect with the underlying dynamic system.

In the third task, we will examine our current autonomous control architectures to determine how machine learning techniques can be integrated with them. Currently, neither 3T nor Remote Agent have built-in learning capabilities. While this project will not provide such "built-in" capability (since we believe that no single learning algorithm is applicable to all of the domains in which these architectures are used), it will produce a design for adding machine learning to autonomous control. In particular, we will design and document interfaces between the machine learning algorithms chosen in Task 2 and our architectures.

We will also explore the option of having two control systems: an off-line system running on simulated data and doing learning; and an on-line system actually controlling the advanced life support machines. When the off-line system has achieved a level of robustness with respect to the simulation, then code is passed to the on-line system for use in control.

### Experimental technique

This is not a thought experiment. We will perform the three tasks listed in the previous section through prototyping and experimentation with actual, working advanced life support systems, with simulations of advanced life support systems and with data gathered from tests of advanced life support systems. Testing of learning algorithms will take place in conjunction with on-going tests of autonomous control algorithms being funded by other sources. There will be considerable leverage off of existing tasks, including the Adjustable Autonomy Testbed, which will provide a discrete event simulation that can be used in our experiments, and the WRS test in December 1999, which will provide data for analysis. We will also use data from early BIO-Plex testing that is expected to begin in FY 2000 and continue in FY 2001.

Prototypes will be implemented in order to gather specific data about how quickly different algorithms can learn control parameters and sequences, what kind of training data they need, whether they can be used in real-time and how robust they are to noisy data. In all cases we will use either actual data or high-fidelity simulations to test algorithms. Quantitative results will be gathered as to the effectiveness of different learning algorithms on this data.

### Related Work

Machine learning is a large and active research area (see [Mitchell, 1997]). We will focus our efforts on machine learning as applied to real-time control systems. This area has received less attention in the machine learning community. While there has been some work in robotic learning, especially using reinforcement learning [Kaelbling et al 1996; Santamaria *et al*, 1998], there has been much less work in applying machine learning to process control. One attempt was a machine learning system to control a heating and power station located in Sweden. In that system, machine learning was used to produce control rules for the injection of NH3 into the combustion chamber [Asker & Bostr, 1995]. It is important to distinguish between machine learning and adaptive control. In the latter, the basic structure of the dynamic model is fixed, leaving only the parameter estimation problem. In our case, the control systems are so experimental that classical adaptive control will not be enough.

Work on integrating machine learning with control architectures has also been scarce, with the notable exception of Soar [Laird et al, 1987]. There have also been recent efforts at doing machine learning off-line and then transferring the results to an on-line control system [Grefenstette et al 1990]. The investigators of this proposal are in close contract those researchers (who are at the Naval Research Lab) in order to learn what can be applied to our problems.

## Status and Milestones

This is a new task. The task leaders have been involved in both autonomous control of advanced life support systems and involved in machine learning research [Boyan & Moore 1998], creating an ideal team for tackling this problem. In particular, the NASA JSC team has been involved in testing of advanced life support autonomous control software during a 15-day test in August 1995 and a 90-day test in the fall of 1997 [Schreckenghost *et al* , 1998a; Schreckenghost *et al* , 1998b]. The same team will also be providing autonomous control software for an upcoming 90 day test of an advanced water recovery system (WRS). In all cases, the 3T architecture [Bonasso *et al* , 1997; Bonasso & Kortenkamp, 1994] is being used. 3T has also been chosen as the baseline for control of BIO-

Plex. The JSC team will be augmented with machine learning specialists from Ames [Boyan & Moore 1998] and Rice University [Subramanian & Gordon, 1996] and by Remote Agent developers at Ames (Gregory Dorais).

### FY 2000 Milestones

Our FY 2000 milestone is to produce a white paper that discusses machine learning issues specific to control of advanced life support systems. This paper will be the result of prototypes and experiments, with testing done using a simulation of the Variable Configuration CO2 Removal (VCCR -- implemented during the Adjustable Autonomy Testbed task in FY 1999) and using data from the WRS test in early FY 2000. The white paper will discuss the trade-offs of different machine learning techniques and how they can be useful in long-term control of advanced life support system. The paper will also discuss integration with Remote Agent and 3T. At the end of FY 2000 we will choose a small number of machine learning techniques and a small number of learning tasks in ALSS and continue working with them in FY 2001.

### FY 2001 Milestones

Our FY 2001 milestone is to prototype the integration of a small number of machine learning techniques (chosen in FY 2000) into both 3T and Remote Agent. This prototype will then be used to control both simulated advanced life support systems and actual tests being conducted at JSC. We will experiment with both on-line and off-line learning, with respect to efficiency and safety. BIO-Plex initial testing is scheduled to begin in FY 2001 and we will begin applying the integrated machine learning and autonomous control system to those tests. We will also assess the usefulness of machine learning to autonomous control of advanced life support systems.

### FY 2002 Milestones

Our FY 2002 milestones include continued experimentation with the learning-enhanced control system in early BIO-Plex tests and accumulation of results. A second milestone is to develop a multi-year plan for building an adaptive control system for the 425-day BIO-Plex test in FY 2006. The multi-year plan will be based on results obtained during three years of research in applying machine learning to the control of Advanced Life Support Systems. We will pursue project (i.e., non-research) funding for the control system of BIO-Plex at this time.

## Customer Relevance

The Human Exploration and Development of Space strategic plan lists research and technology development for advanced life support systems as an explicit goal of the Enterprise. Any advanced life support system, especially if a long-duration mission is envisioned, will need intelligent autonomous control that can adapt to changing circumstances and requirements. The BIO-Plex project is a major HEDS initiative to test and demonstrated advanced life support systems here on earth. This research will directly support these efforts. We have had numerous discussions with BIO-Plex engineers (Terry Tri, Mary Beth Edeen, Karen Meyers) about their autonomous control requirements and feel that this proposal addresses several of them. JSC/EC has already committed money to fund research into autonomous control of advanced life support systems, which will complement this project.

The Space Science strategic plan lists research into autonomous operations and intelligent systems as an explicit goal of the Enterprise. In particular, the in-situ propellant production and the robotics initiatives in Space Science will benefit directly from this research. We expect this research to also support adaptive on-board autonomy, which is a cross-enterprise requirement.

Since this is a pull task, we have enclosed a letter of support from Terry Tri, program manager of BIO-Plex at NASA Johnson Space Center.

## Technical References

[Asker & Bostr, 1995] Lars Asker and Henrik Bostr, "Building the DenNOx System: Experiences from a Real-World Application of Machine Learning," Technical Report 95-023, Department of Computer and Systems Sciences (DSV), Stockholm University, 1995.

[Bonasso & Kortenkamp, 1994] R. Peter Bonasso and David Kortenkamp, "An Intelligent Agent Architecture in Which to Pursue Robot Learning," *Working Notes: MCL/COLT '94 Robot Learning Workshop*, 1994.

[Bonasso *et al*, 1997] R. Peter Bonasso, R. James Firby, Erann Gat, David Kortenkamp, David Miller and Marc Slack, "Experiences with an Architecture for Intelligent, Reactive Agents," *Journal of Experimental and Theoretical Artificial Intelligence*, Vol. 9, No. 1, 1997.

[Boyan & Moore 1998] Boyan, J. A. and A. W. Moore, "Learning Evaluation Functions for Global Optimization and Boolean Satisfiability," *Fifteenth National Conference on Artificial Intelligence (AAAI*), 1998.

[Grefenstette *et al*, 1990] John Grefenstette, C. L. Ramsey and Alan C. Schultz, "Learning sequential decision rules using simulation models and competition," *Machine Learning*, 5(4):355-381, 1990.

[Kaelbling et al, 1996] Leslie Pack Kaelbling, M. L. Littman and A. W. Moore, "Reinforcement Learning: A Survey," *Journal of Artificial Intelligence Research*, Vol. 4, pages 237-285, 1996.

[Laird et al, 1986] John E. Laird, Allen Newell, and Paul S. Rosenbloom, "Soar: An Architecture for General Intelligence," *Artificial Intelligence*, 33(1), 1987.

[Mitchell, 1997] Thomas Mitchell, *Machine Learning*, McGraw Hill, New York, 1997.

[Moore et al, 1998], Moore, A. W., J. G. Schneider, J. A. Boyan and M. S. Lee. "Q2: Memory-Based Active Learning for Optimizing Noisy Continuous Functions," *Machine Learning: Proceedings of the Fifteenth International Conference (ICML)*, 1998.

[Santamaria et al, 1998] Juan C. Santamaria, Richard C. Sutton, and Ashwin Ram, "Experiments with Reinforcement Learning in Problems with Continuous State and Action Spaces," *Adaptive Behavior*, 6(2):163-217, 1998.

[Schreckenghost *et al*, 1998a] Debra Schreckenghost, Daniel Ryan, Carroll Thronesbery, R. Peter Bonasso and Daniel Poirot, "Intelligent Control of Life Support Systems for Space Habitats," *Proceedings of the Conference on Innovative Applications of Artificial Intelligence*, 1140--1145, 1998.

[Schreckenghost *et al*, 1998b] Debra Schreckenghost, Mary Beth Edeen, R. Peter Bonasso and Jon Erickson, "Intelligent Control of Product Gas Transfer for Air Revitalization," *Proceedings of 28th International Conference on Environmental Systems*, 1998.

[Subramanian & Hunter, 1992] Devika Subramanian and S. Hunter, "Estimating utility and the design of provably good EBL algorithms," *Proceedings of the 9th International Conference on Machine Learning*, pp. 426-435, 1992.

[Subramanian & Gordon, 1993] Devika Subramanian and Diana Gordon, "A multi-strategy learning scheme for knowledge assimilation in embedded agents," *Informatica*, 17:8-15, 1993.

[Subramanian & Gordon, 1996] Devika Subramanian and Diana Gordon, "Comparison of Action Selection Learning Methods," *Proceedings of the 3rd International Workshop on Multi-stretegy Learning*, pp. 95-102, 1996.